

Interpretable Digital Twins for optimizing Alzheimer's disease drug trial sample sizes for sex-specific efficacy analyses

Daniel Andrews, PhD^{1,2}, D. Louis Collins, PhD^{1,2,3}

¹Department of Biomedical Engineering, McGill University, Montreal, Canada. ²McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada.

³Department of Neurology & Neurosurgery, McGill University, Montreal, Canada.



This work was supported by an
AAIC 2026 Conference Fellowship



Open-access methods paper in
Alzheimer's Research & Therapy
DOI: 10.1186/s13195-026-02024-5
BMC · Springer Nature

Contact: daniel.andrews2@mail.mcgill.ca

BACKGROUND

- Lecanemab's Phase 3 trial CLARITY AD indicated a sex difference in the drug's efficacy.^{1,2}
- Future trials may evaluate sex-specific efficacy, requiring adequately powered sex subgroups.
- In simulations of CLARITY AD, we combine interpretable Digital Twins with prediction-powered inference (PPI) to reduce sex subgroup sample sizes vs. a standard analysis.³

METHOD

1) Modelled cognitive decline:

- Linear mixed effects models tracked 5-year CDRSB change from baseline (CDRSB Δ bl) separately in 316 male and 209 female ADNI subjects meeting CLARITY AD criteria.

2) Simulated sex-specific trial data:

- Resampled subjects from approximately half of each sex dataset. Recruits were randomized 1:1 to drug and placebo groups. CDRSB Δ bl scores at 18 months were interpolated using the sex-specific models, encoding a 25% CDRSB Δ bl-reducing drug effect.

3) Estimated drug effect:

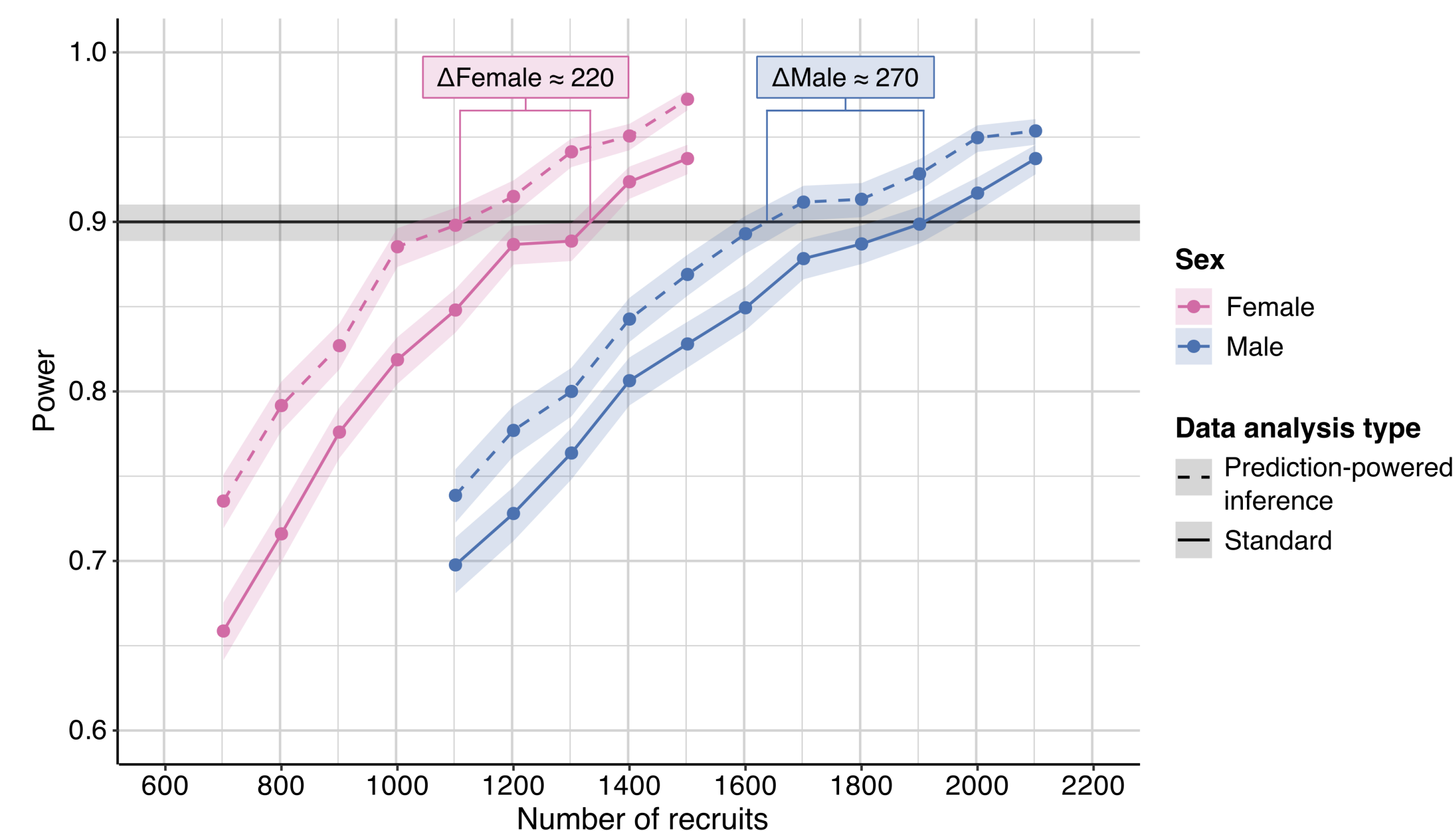
- Standard analysis: Used a *t*-test to estimate drug vs. placebo group difference in mean 18-month CDRSB Δ bl.
- PPI: Predicted each recruit's CDRSB Δ bl as the mean of their most-similar Digital Twins, ranked by Gower's distance between the recruit's baseline demographic-clinical profile and the profiles of subjects in the independent half of the sex-matched dataset. Used a *t*-test to estimate drug vs. placebo difference in mean prediction-adjusted CDRSB Δ bl.

4) Repeated simulations & calculated power:

- Power is the proportion of 3,000 simulations with a statistically significant drug effect estimate.

RESULTS

Figure 1. Sex-specific power curves ($\pm 95\%$ CIs).



- PPI required $\sim 16.5\%$ fewer female and $\sim 14.1\%$ fewer male recruits to reach 0.9 power vs. the standard analysis (Fig. 1).
- Predicted and known CDRSB Δ bl correlated more highly when using most-similar Twins than random Twins (Figs. 2 and 3).
- Sex-specific optimal numbers of Twins per recruit were used (Figs. 2 and 3).

Figure 2. Correlation of known and predicted female CDRSB Δ bl ($\pm 95\%$ CIs).

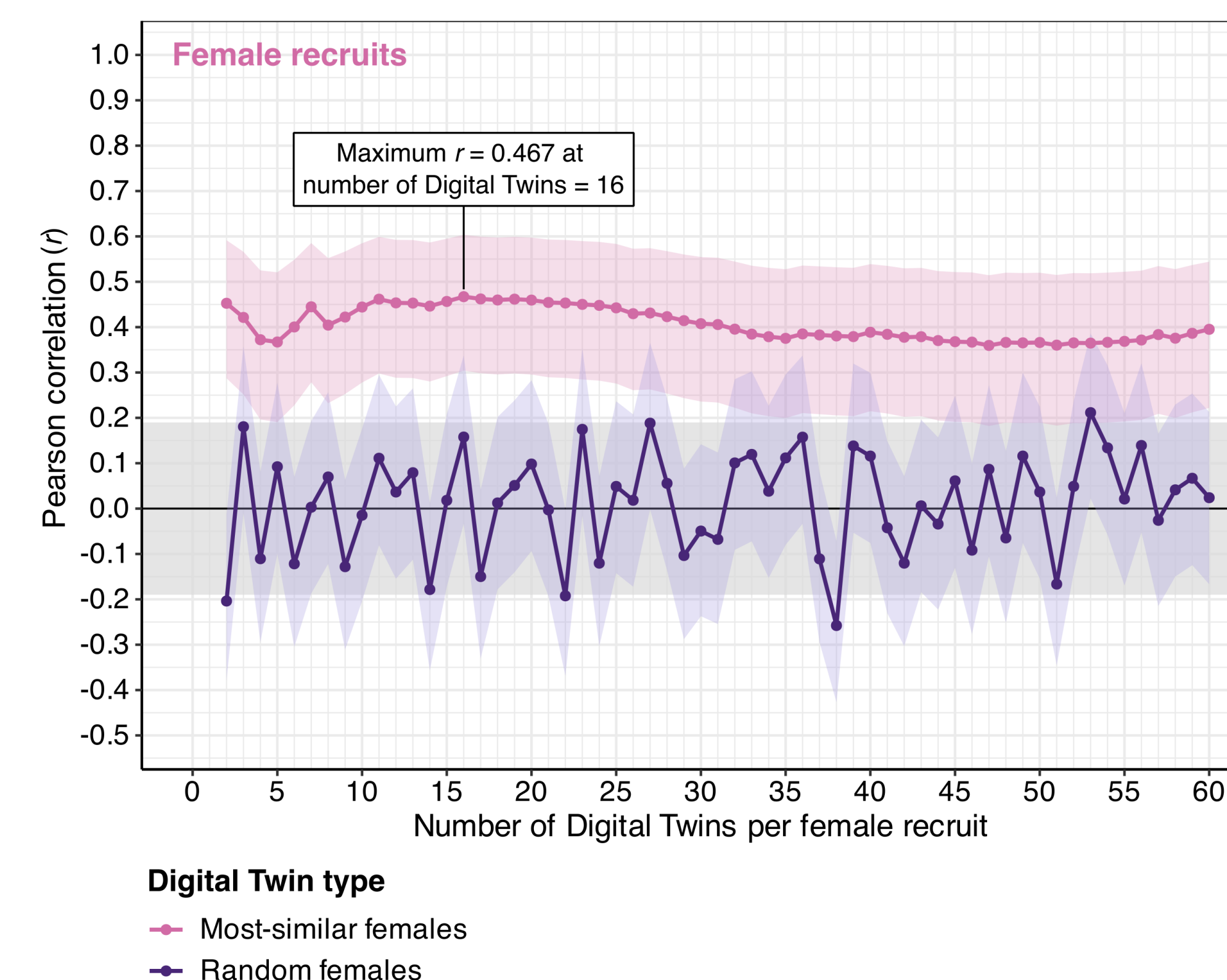
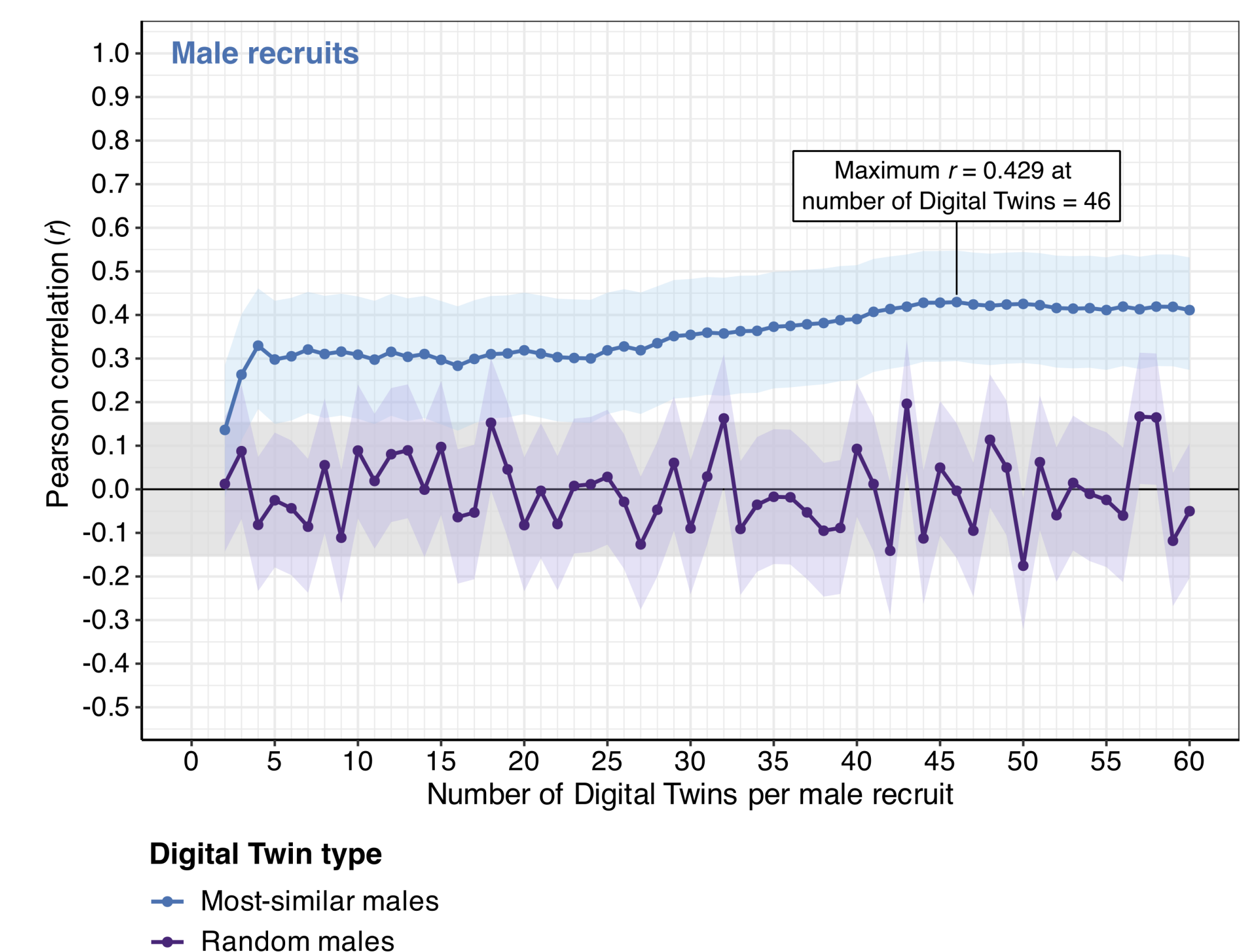


Figure 3. Correlation of known and predicted male CDRSB Δ bl ($\pm 95\%$ CIs).



CONCLUSION

- Prediction-powered inference with interpretable Digital Twins reduces required sample sizes for sex-specific Alzheimer's drug efficacy analyses.

REFERENCES

- 1) Eisai & Biogen (2019). clinicaltrials.gov/ct2/show/NCT03887455
- 2) van Dyck et al. (2022). *NEJM*. doi.org/10.1056/NEJMoa2212948
- 3) Andrews, Golchi, Collins (2026). *Alz. Res. Ther.* doi.org/10.1186/s13195-026-02024-5